



EGUsphere, referee comment RC3
<https://doi.org/10.5194/egusphere-2022-98-RC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-98

Anonymous Referee #3

Referee comment on "Constraining low-frequency variability in climate projections to predict climate on decadal to multi-decadal timescales – a poor man's initialized prediction system" by Rashed Mahmood et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-98-RC3>, 2022

General Comments:

This manuscript by Mahmood et al. (egusphere-2022-98) explores an alternate way of decadal to multidecadal predictions by subsampling individual CMIP6 historical simulations that better matches observed SST at a given time (start dates) and by tracking the trajectories of the subsampled simulations for the next two decades (analogue method hereafter). The authors show that the added value of the analogue method over the uninitialized simulations is comparable to that of initialized decadal prediction simulation (DCPP). The manuscript is overall well organized and written. I have enjoyed reading the manuscript. I also acknowledge the thorough effort to test the sensitivity of the results to subsampling criteria. I think the manuscript has potential to draw attention from climate research community, as the proposed method can leverage existing simulations in the application of the prediction of climate, instead of running computationally expensive decadal prediction simulations. However, before I recommend accepting for publication, I have a couple of specific comments that I wish the authors further demonstrate or explain along with some minor suggestions.

Specific Comments:

1) The authors shows that the analogue method exhibits high skill in the Pacific Ocean, even higher than skill in the subpolar North Atlantic, which even lasts for FY11-20 . Such a long memory in the Pacific is surprising and in stark contrast to the current understanding that predictability in the Pacific Ocean is low on decadal timescales while very high in the subpolar North Atlantic. The low predictability in initialized decadal predictions may be related to initialization shock/drift, as the authors also discuss in the manuscript. However, the low predictability in the Pacific Ocean is also pervasive in “perfect model” experiments (e.g., Collins 2002; Pohlmann et al. 2004), which does not suffer from initialization shock/drift. Why the author’s analogue method shows such superior skill in the Pacific Ocean? Isn’t this high skill possibly related to the forced signal that is not completely removed by the method the authors used (Smith et al. 2019)? One way to verify this is to perform a bootstrapping method for the statistical test, rather than Student’s t-test. If the ACC from Best30 is found outside of the (eg., 2.5 to 97.5 percentile) distribution of the ACCs from randomly sampled 30 members, assuming that the uninitialized ensemble mean used in Smith et al.’s method is the total 212-member ensemble mean, the authors can say more confidently that the high Pacific skill is indeed not from the forced signal.

2) If that is the case, why the skill is so high in the Pacific? Since this would be the most important finding of the study, in my opinion, as it is in contrast to the current understanding, I recommend that the authors further demonstrate the reasons for the high Pacific skill.

3) The Atlantic skill is low for FY1-10, but picks up for FY11-20 (Fig. 2d-e). Why is this the case? I think the low skill for FY1-10 is because Best30 is dominated by the correlations in the Pacific and as demonstrated in the regional SST constraints, but it is hard to understand why there is a rebound in ACC skill.

4) The authors introduce several statistical methods in section 2, without a description, just by referring to citations. I recommend adding a brief description for each method.

Technical corrections:

l. 43: Remove "in" after phasing.

l. 88: ...anomalies "relative to" the reference climatological period...